

データ利活用に対する 人々の評価と日本の未来

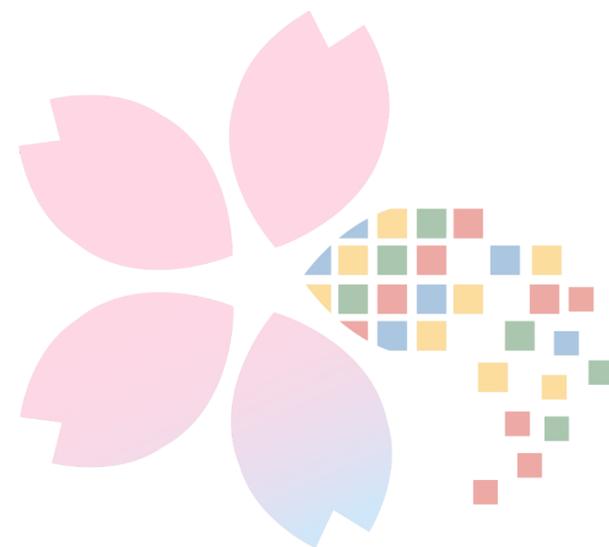
2019.7.24

山口真一（国際大学GLOCOM 講師・主任研究員）

Thanks to

佐相弘明（国際大学GLOCOM リサーチ・アソシエイト）

青木志保子（国際大学GLOCOM 主任研究員）



プロフィール



国際大学グローバル・コミュニケーション・センター 講師/主任研究員
東洋英和女学院大学 非常勤講師
グリー株式会社 アドバイザリーボード

Recent Key Interest: ソーシャルメディア / プラットフォーム / データ活用

2010年慶應義塾大学経済学部卒、2015年同大学経済学研究科で博士号（経済学）を取得し、国際大学グローバル・コミュニケーション・センター助教を経て、2016年より現職。専門は計量経済学。研究分野は、ネットメディア論、フリー・ビジネス、データ活用戦略等。「あさイチ」「ニュースウォッチ9」（NHK）や「日本経済新聞」をはじめとして、メディアにも多数出演・掲載。組織学会高宮賞受賞（2017年）、情報通信学会論文賞受賞（2017年・2018年）、電気通信普及財団賞受賞（2018年）。主な著作に『炎上とクチコミの経済学』（朝日新聞出版）、『ネット炎上の研究』（勁草書房）、『ソーシャルゲームのビジネスモデル』（勁草書房）などがある。



【計量経済学】

理論に基づいて数学的モデルを構築し、統計学的手法で定量的な実証分析を行う学問。GDP算出などに利用。

- ① 要因が特定できる
- ② 効果を定量的に把握できる
- ③ 将来予測をもとに最適な戦略を導出できる

はじめに：
プラットフォーム時代とデータ利活用

はじめに：プラットフォーム時代とデータ利活用

◆ プラットフォームとデータ利活用

- **プラットフォームの21世紀**（公文, 2015）と言っても過言ではないほど、現代の人々の生活にはプラットフォーム（特にB to C）が入り込んでいる。
- ①ネットワーク効果、②**データ利活用**の2つにより巨大化（中田, 2015）。

◆ データ利活用に関する議論が国内外で活発化

- 自民党経済成長戦略本部「『データ』を利活用できる環境をいち早く整備する」。
- Society5.0。



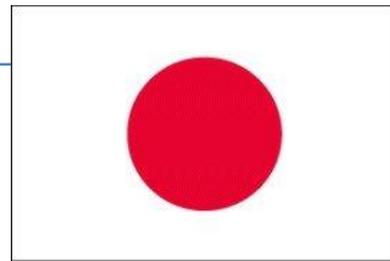
はじめに：プラットフォーム時代とデータ利活用

◆ データ利活用に関する議論の現状



自由なデータ利活用

https://www.lifehacker.jp/2014/11/141122company_culture.html



規制・人権

<https://www.hellocybernetics.tech/entry/2017/02/03/063338>



他国の動き

**消費者の「利便性」と「不利益」に着目し
その意向と影響を実証的に論じているものは少ない**

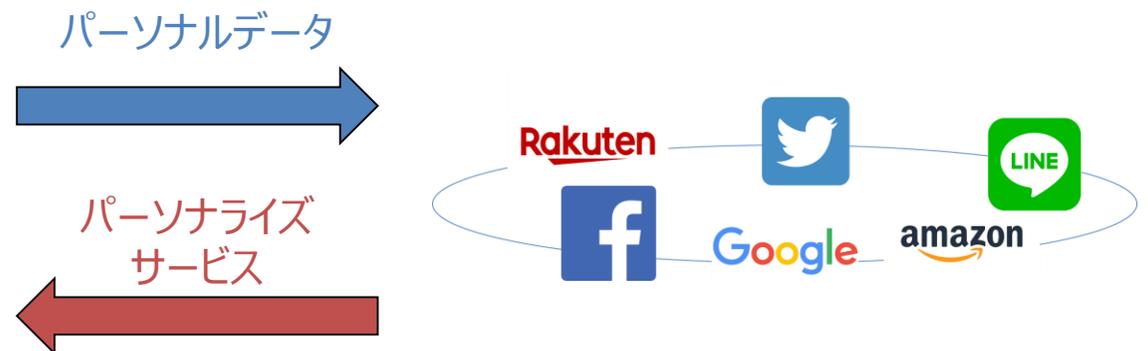
パーソナルデータの利活用については、どうしてもプライバシー保護の観点からの法制度議論が中心となっている。もちろんそれも重要だが、同時に経済学的検討も今後必要ではないか。

——高崎・高口（2015）

はじめに：プラットフォーム時代とデータ利活用

◆ 国内実証研究は少ない

- 少ない中でも、多くが「パーソナルデータ提供への評価」を測ったり、「不安感」を調査したりするもの（三菱総合研究所, 2017など）。海外も同様。
- データ提供が「コスト」で、パーソナライズされたサービスを「対価」と考えるならば、それを消費者目線で総合的に実証分析しないと、社会的厚生にとって適切な制度は検討できない。



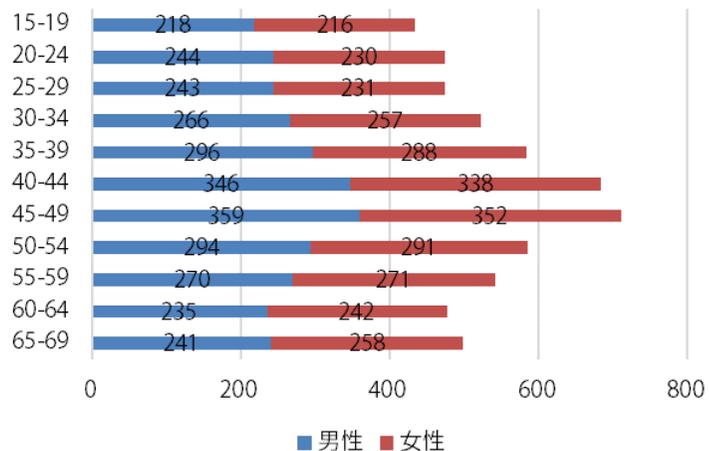
本研究では、データ利活用に対する総合的な評価を実証的に分析し、消費者にとって適切なデータ利活用の在り方を検討する。

はじめに：プラットフォーム時代とデータ利活用

◆ 分析方法

アンケート調査分析

- 実施期間：2019年1月。
- 対象：ネットリサーチ会社マイボイスコム社の15歳～69歳の登録モニタ。
- 性年代別のインターネットユーザ数に応じて層化抽出法で取得。
- サンプルサイズ**5,986人**。



ヒアリング調査

- **20代の若者10名**に対して、データ利活用への考え方などヒアリング調査。
- 職業・年齢がばらけるように取得。

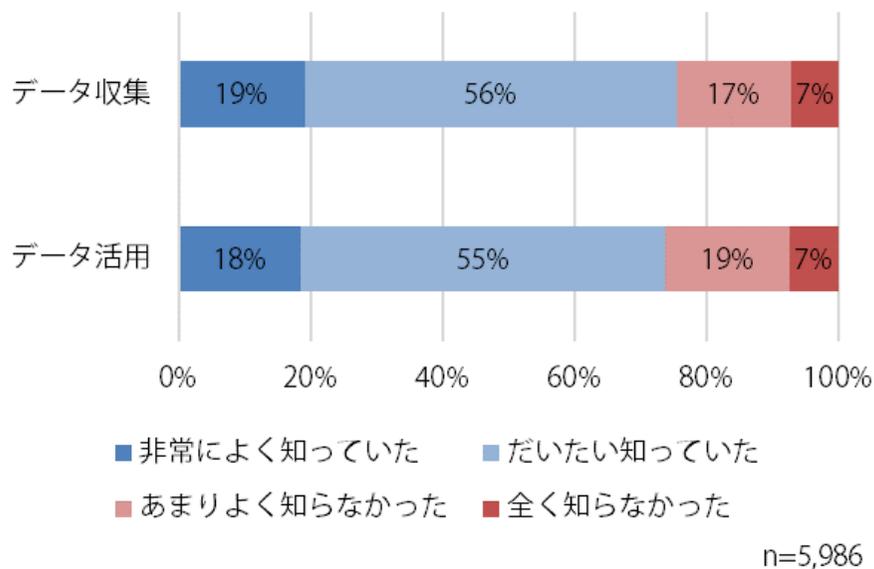


データ利活用に対する不安と利便性

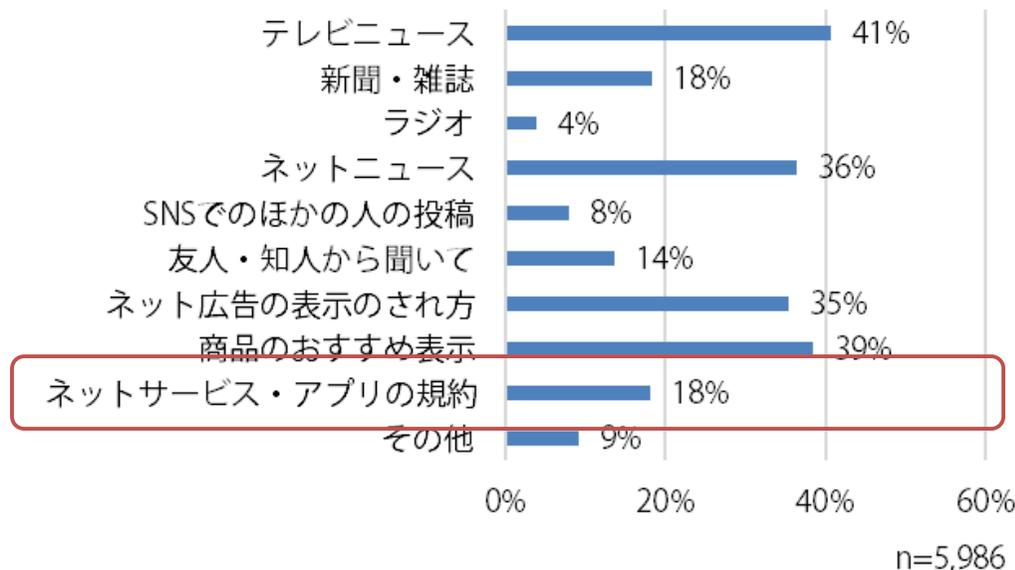
データ利活用に対する不安と利便性

◆ データ利活用の認知度

- データ利活用を認知している人は約75%。年代による傾向なし。
- 認知経路として、「規約」は約18%に留まる。



データの利活用に対する認知割合



データの利活用を何で知ったか
(複数回答／認知者が対象)

データ利活用に対する不安と利便性

◆ 規約の例

- 2.1 お客様は、本規約の定めに従って本サービスを利用しなければなりません。お客様は、本規約に有効かつ取消不能な同意をしないかぎり本サービスを利用できません。
- 2.2 お客様が未成年者である場合は、親権者など法定代理人の同意を得たうえで本サービスを利用してください。また、お客様が本サービスを事業者のために利用する場合は、当該事業者も本規約に同意したうえで本サービスを利用してください。
- 2.3 お客様は、本サービスを実際に利用することによって本規約に有効かつ取消不能な同意をしたものとみなされます。
- 2.4 本サービスにおいて個別利用規約がある場合、お客様は、本規約のほか個別利用規約の定めにも従って本サービスを利用しなければなりません。

3. 規約の変更

当社は、当社が必要と判断する場合、あらかじめお客様に通知することなく、いつでも、本規約および個別利用規約を変更できるものとします。変更後の本規約および個別利用規約は、当社が運営するウェブサイト内の適宜の場所に掲示された時点からその効力を生じるものとし、お客様は本規約および個別利用規約の変更後も本サービスを使い続けることにより、変更後の本規約および適用のある個別利用規約に対する有効かつ取消不能な同意をしたものとみなされます。かかる変更の内容をお客様に個別に通知することはいたしかねますので、本サービスをご利用の際には、随時、最新の本規約および適用のある個別利用規約をご参照ください。

4. アカウント

- 4.1 お客様は、本サービスの利用に際してお客様ご自身に関する情報を登録する場合、真実、正確かつ完全な情報を提供しなければならず、常に最新の情報となるよう修正しなければなりません。
- 4.2 お客様は、本サービスの利用に際してパスワードを登録する場合、これを不正に利用されないようご自身の責任で厳重に管理しなければなりません。当社は、登録されたパスワードを利用して行なわれた一切の行為を、お客様ご本人の行為とみなすことができます。
- 4.3 本サービスに登録したお客様は、いつでもアカウントを削除して退会することができます。
- 4.4 当社は、お客様が本規約に違反または違反するおそれがあると認めた場合、あらかじめお客様に通知することなく、アカウントを停止または削除することができます。
- 4.5 当社は、最終のアクセスから1年間以上経過しているアカウントを、あらかじめお客様に通知することなく削除することができます。

Amazonからのメールや、ソーシャルメディア画面上での広告で、自分が好きなものがおすすめされることを通じて気づいた。

——22歳女性

データ収集についてはテレビのニュースで聞いて知った。

——25歳男性

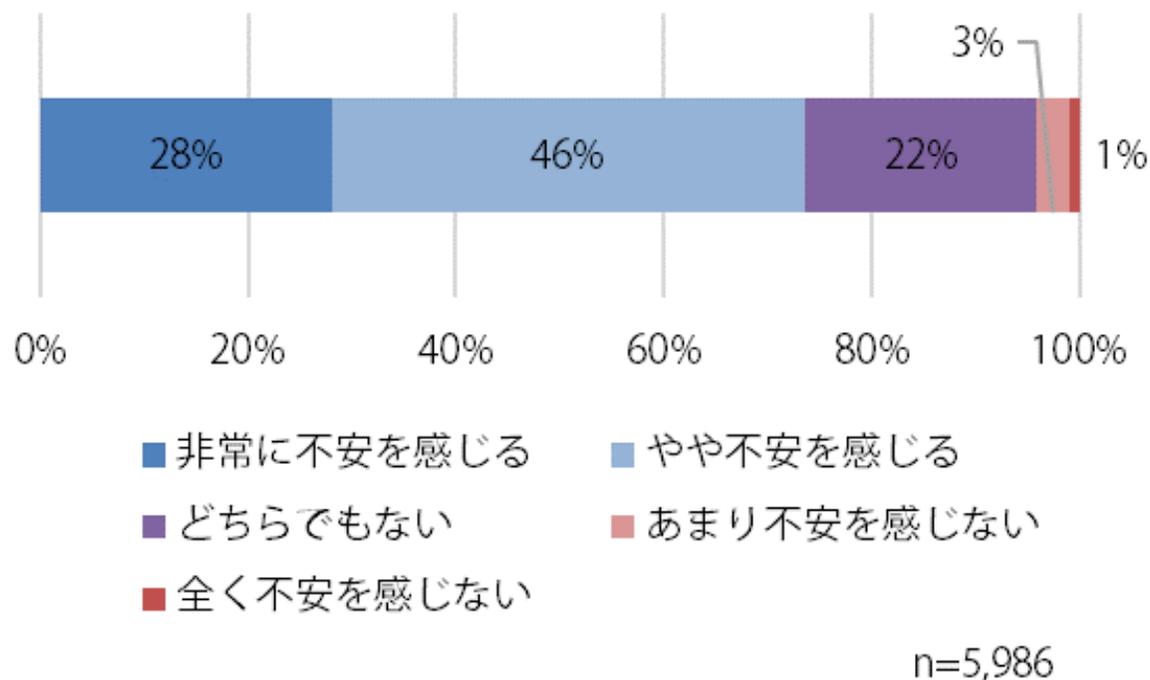
データが収集・分析されていることは知っていた。その理由は、普段サービスを利用している中で、「おすすめ」や「閲覧履歴からみる」というものが明らかに自分にあったものであり、わかった。尚、規約は細かいところまで特に読んでいない。

——28歳男性

データ利活用に対する不安と利便性

◆ データ利活用に不安を感じる人

➤ データ利活用に対して不安に感じている人は**74%**いる。



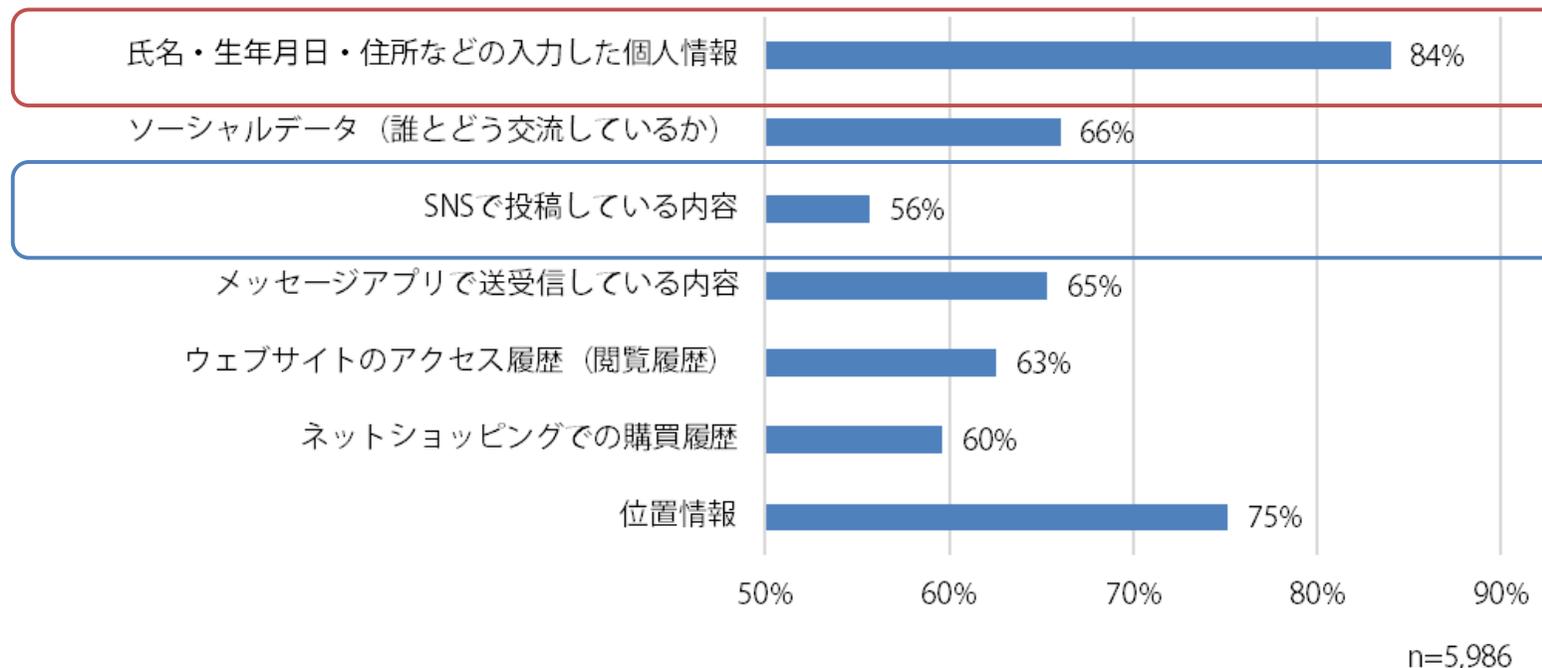
- 不安な人が多いというのは三菱総合研究所（2017）の調査結果と同様。
- 同研究では、パーソナルデータ提供に対して日本、韓国、中国といった**アジア圏の国で不安感が強い**ことが示されている。

データ利活用を不安に感じるか

データ利活用に対する不安と利便性

◆ 特に不安なのは「入力した個人情報」

- 「氏名・生年月日・住所」に対しては特に不安感が高い。
- 「SNSでの投稿内容」はプライベートだが不安感は低い。「リアルな自分と直接つながるデータ」が不安。

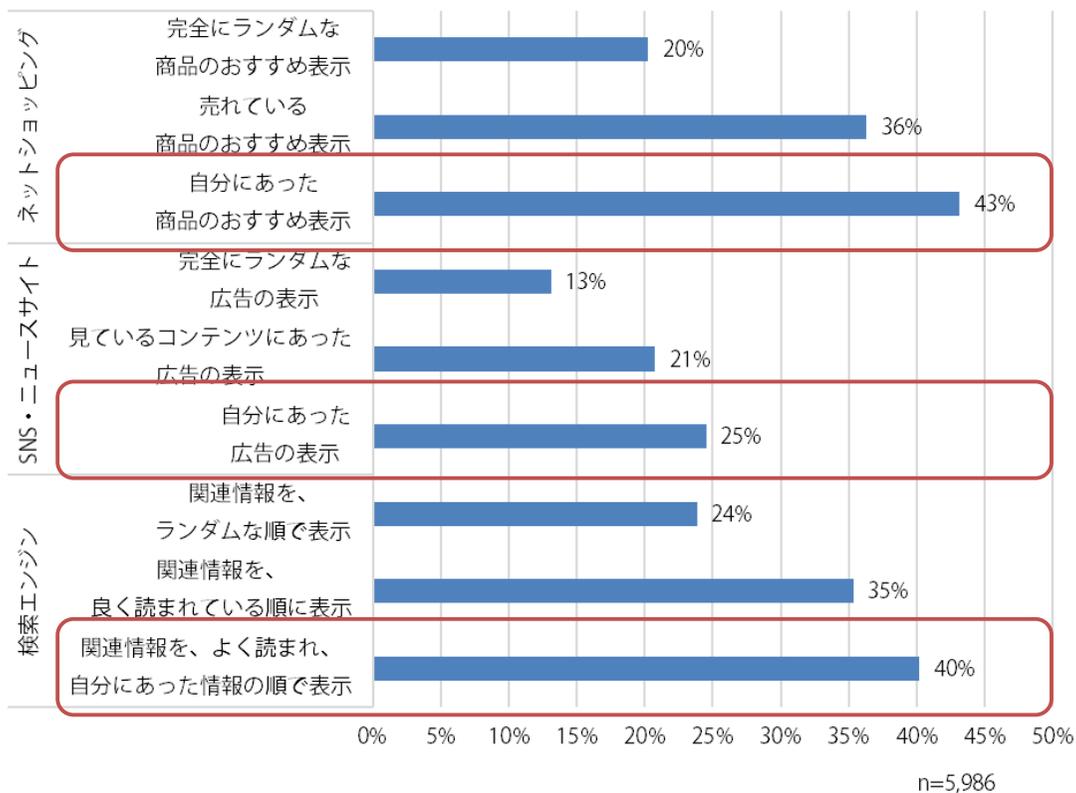


各データ利活用に不安な人の割合

データ利活用に対する不安と利便性

◆ データ利活用に利便性を感じる人

- 不安な人が多い一方、パーソナルデータを活用した機能を評価する人も多い。



データの利活用がもたらす機能への評価
(あった方が良いと思う人の割合)

- 以下の3点について、あった方が良いと思う人の割合を調査。
 - ① データの利活用を一切しない
 - ② マクロ的にデータを利活用する
 - ③ 個人レベルでデータを利活用する

※「非常にあった方が良いと思う」～「まったくいらないと思う」の5段階で評価してもらい、「非常にあった方が良いと思う」「だいたいあった方が良いと思う」のいずれかを選択した人の割合。

データ利活用への評価の経済的測定

データ利活用への評価の経済的測定

◆ データ利活用への評価を経済的に測る

- 「**支払い意思額**」の推定をとおして、評価を定量的に測る。
- 仮想評価法（**CVM** : Contingent Valuation Method）をベースに設問を設計。

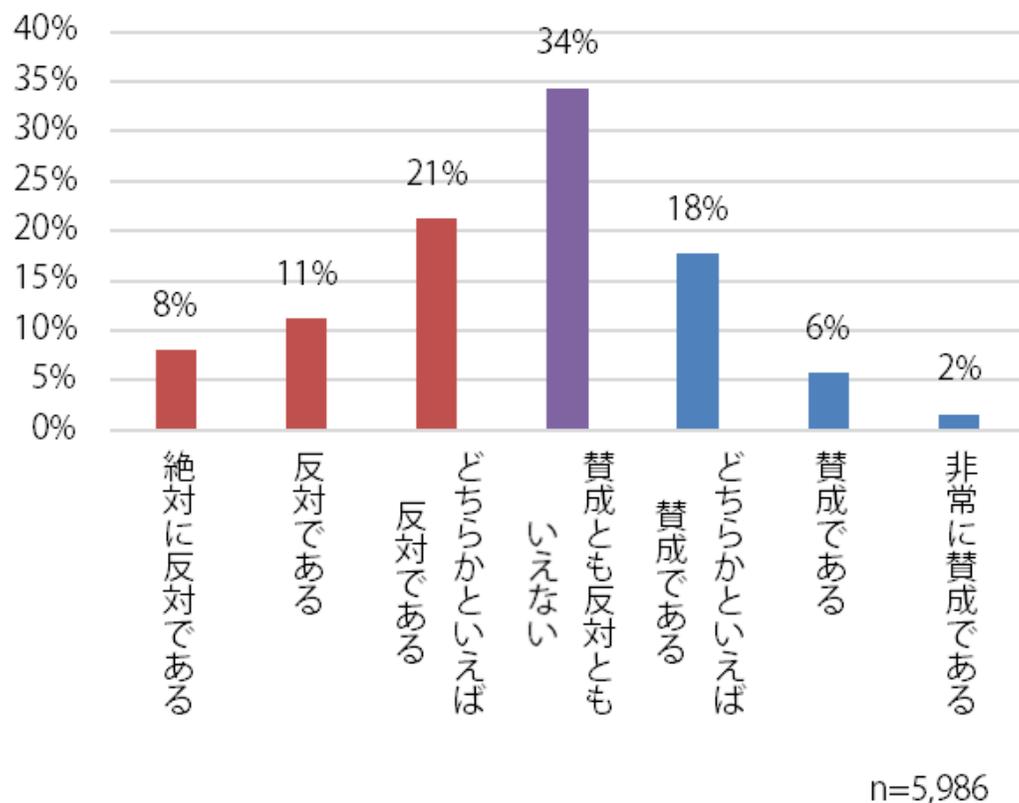
- CVMとは、環境を守る、あるいは実現するために支払っても構わない最大の金額を尋ねることによって、その環境の価値を金額として定量的に評価する手法。
- 環境分野など、**市場の存在しない財・サービス**の評価に幅広く利用（栗山, 1997）。

◆ 研究・裁判などで幅広く使われるCVM

- 高口・実積（2013）：携帯電話の各要素に対する評価（重視度）を分析。
- 行平ほか（2018）：イベントにおける帆船誘致の評価額を推計。
- オハイオ裁判（1989）：CVMを自然資源損害額の算出方法として確認。
- 石油汚濁防止法（1990）：被害額算出法と認定。

データ利活用への評価の経済的測定

◆ データ利活用に対する総合評価



- データ利活用に対しては、「お金を払ってでもやってほしい人」と、「お金を払ってでもやめてほしい人」が存在すると考えられる。
- 分析前に、ネットサービス全般におけるデータ利活用への総合的評価を調査。
- **賛成の人26%**をデータ利活用にポジティブな人とする。
- **反対の人40%**をデータ利活用にネガティブな人とする。

ネットサービス全般におけるデータ利活用に対する総合評価

データ利活用への評価の経済的測定

◆ 分析対象とするサービス

- 8つのサービスカテゴリの中で、特に利用者の多い**主要20サービス**。
- 各具体的なサービスの、「データ利活用に対する評価」を経済的に測る。

- SNS (Facebook・Twitter・Instagram・TikTok)
- メッセージアプリ (LINE・Messenger)
- ネットニュース (Yahoo!ニュース・LINE NEWS)
- 検索エンジン (Google・Yahoo!検索)
- クチコミサイト (食べログ・@cosme・楽天トラベル)
- 動画サイト (YouTube・ニコニコ動画)
- フリマアプリ・オークションサイト (メルカリ・ラクマ・ヤフオク!)
- ショッピングサイト (Amazon・楽天)

データ利活用への評価の経済的測定

◆ 支払い意思額取得方法①

- ▶ **利用しているサービスすべて**について「パターン1」「パターン2」のいずれかを提示し、**それぞれのサービス**におけるデータ利活用に対する支払意思額を取得。

◆Q25でデータの収集と活用に「非常に賛成である」「賛成である」「どちらかといえば賛成である」を選んだ方にお伺いします。

「無料で利用している」と前半の質問でお答えいただいた以下のサービスが、仮にあなたのデータを収集・活用しなくなったと想定してください。

※あなたのデータを収集・活用しないと、あなたにあった(個人個人にカスタマイズした)サービスや広告の提供は止められますが、個人情報流出や悪用のリスクは低下します。

※あなたにあった(個人個人にカスタマイズした)サービスや広告の提供とは以下のようなことを指します。

- ・あなたにあった商品のおすすめ表示(ネットショッピングにおける)
- ・あなたにあった広告の表示(SNS、ニュースサイトなどにおける)
- ・検索したワードに関連した情報について、あなたにあった情報の順で表示(検索エンジンにおける)
- ・あなたにあった動画のおすすめ表示(動画サイトにおける)

Q26. お金を支払えば、これまで通り以下のサービスがデータを収集・活用するようになるとしたら、あなたはひと月あたりいくらまで支払っても良いと思いますか。その支払ってよい最高額(税込み)について、最も近いものを1つお選びください。

※ただし、データを収集・活用されないまま無料で利用する場合は、0円をお選びください。

パターン1：第1段階で「非常に賛成である」「賛成である」「どちらかといえば賛成である」だった人向け

データ利活用への評価の経済的測定

◆ 支払い意思額取得方法②

◆ Q25でデータの収集と活用に「絶対に反対である」「反対である」「どちらかといえば反対である」を選んだ項目の、各サービスについてお伺いします。

「無料で利用している」と前半の質問でお答えいただいた以下のサービスが、仮にあなたのデータを収集・活用しないプランも提供したと想定してください。

※あなたのデータを収集・活用しないと、あなたにあった(個人個人にカスタマイズした)サービスや広告の提供は止められますが、個人情報流出や悪用のリスクは低下します。

※あなたにあった(個人個人にカスタマイズした)サービスや広告の提供とは以下のようなことを指します。

- ・あなたにあった商品のおすすめ表示(ネットショッピングにおける)
- ・あなたにあった広告の表示(SNS、ニュースサイトなどにおける)
- ・検索したワードに関連した情報について、あなたにあった情報の順で表示(検索エンジンにおける)
- ・あなたにあった動画のおすすめ表示(動画サイトにおける)

Q27. お金を支払えば、そのデータを収集・活用しないそのプランを利用できるとすれば、あなたはひと月あたりいくらまで支払っても良いと思いますか。

その支払ってよい最高額(税込み)について、最も近いものを1つお選びください。

※ただし、データを収集・活用されるまま無料で利用する場合は、0円をお選びください。

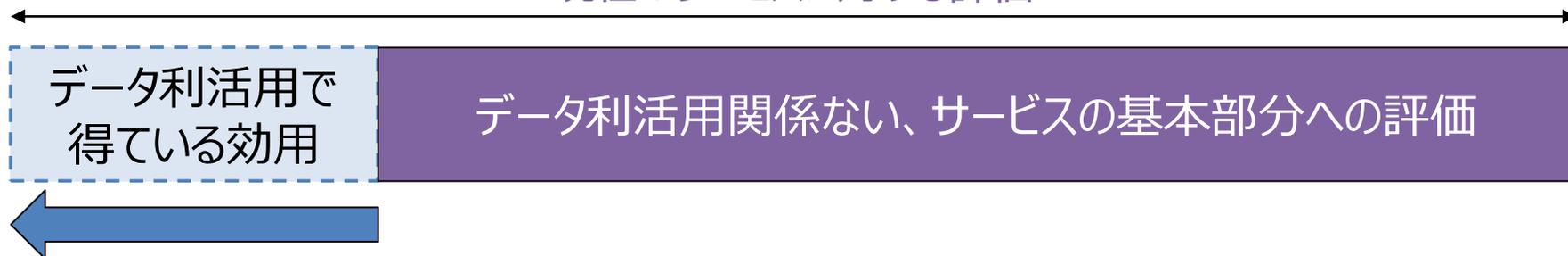
パターン2：第1段階で「絶対に反対である」「反対である」「どちらかといえば反対である」だった人向け

データ利活用への評価の経済的測定

◆ 推計する部分

1. データ利活用にポジティブな人（パターン1）

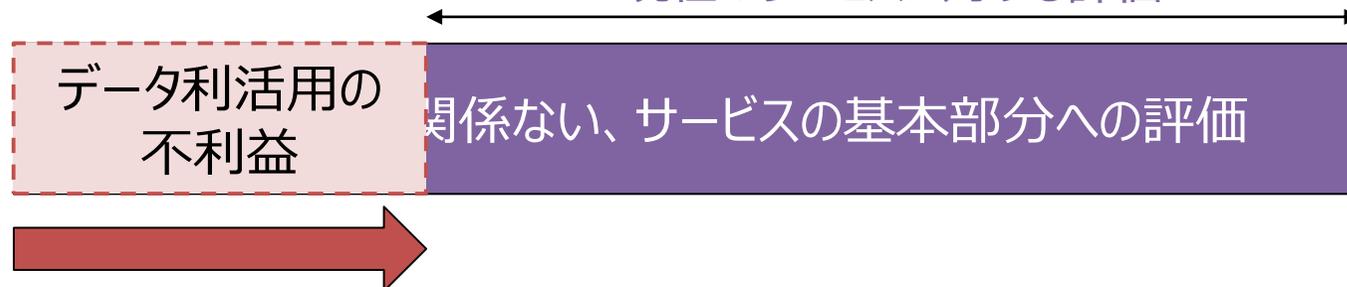
現在のサービスに対する評価



支払い意思額はプラスとなる

2. データ利活用にネガティブな人（パターン2）

現在のサービスに対する評価



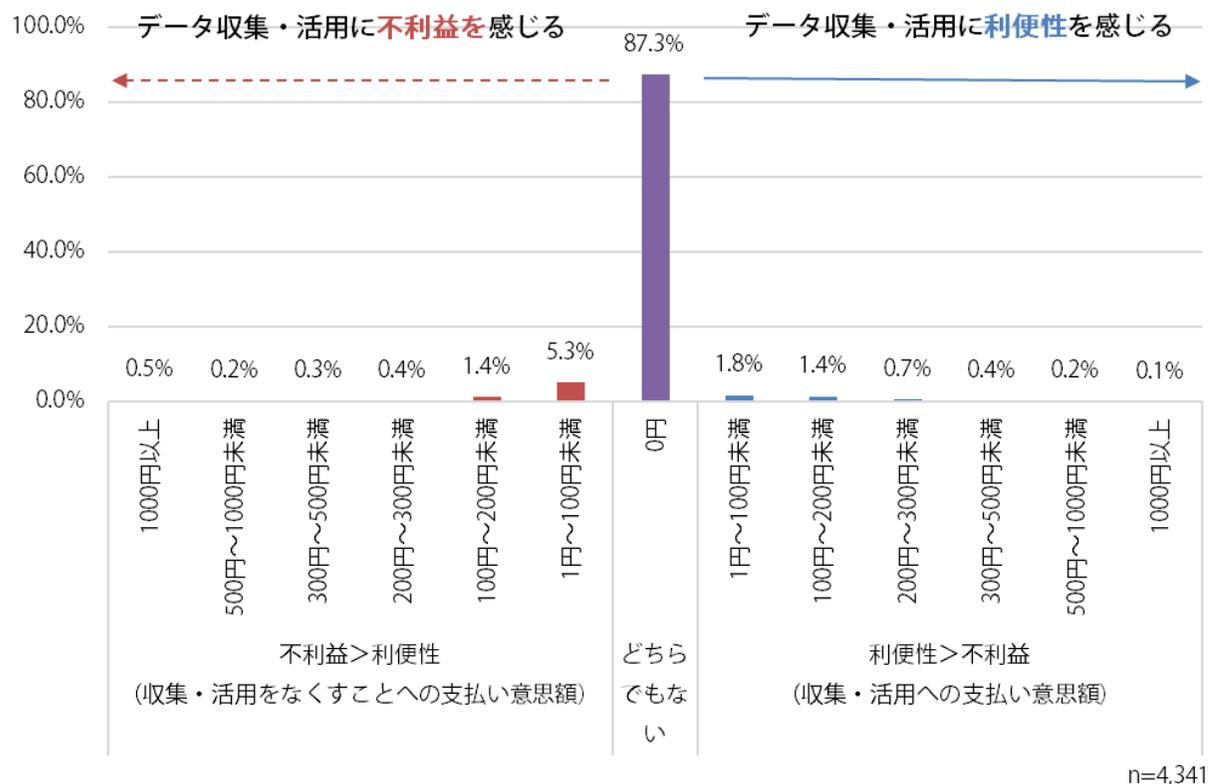
支払い意思額はマイナスとなる

データ利活用への人々の評価

データ利活用への人々の評価

◆ 支払い意思額の分布

➤ 全てのサービスについて**85%~97%**の利用者は支払い意思額0円であった。

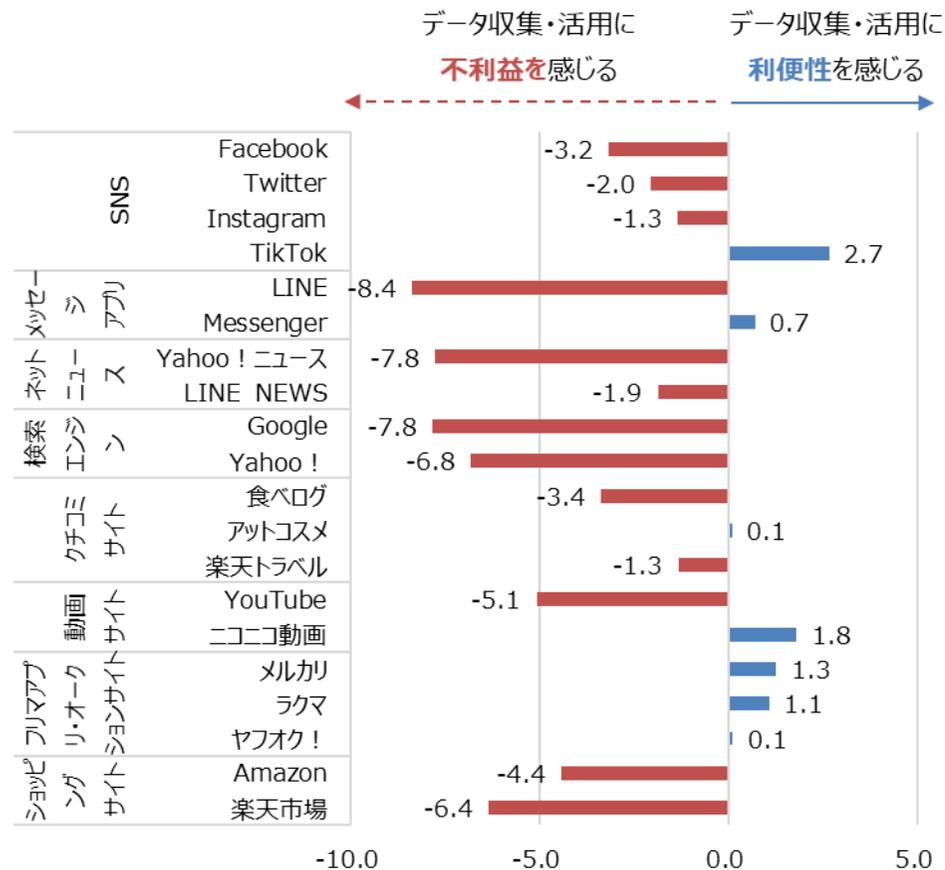


- 多くの人にとって「不安に感じる」と「利便性を感じる」効果は打ち消し合っている。
- **支払ってまで**データの利活用をなくしたり、あるいはこのまま続けて欲しいと強く希望したりする人はほとんどない。

利用者の最も多かったYouTubeにおける支払い意思額分布

データ利活用への人々の評価

◆ データ利活用に対する平均支払い意思額（月額）



データ利活用に対する支払い意思額平均値（円／月額／利用者のみ）

- 支払い意思額の平均が**マイナス**となっているサービスが大半を占めている。
- 総じて支払い意思額の絶対値は小さい。最も大きかったのはLINEであり、月額－8.4円となった。
- 主たる利用方法が**オープンかどうか**が、そのサービスのデータ利活用に対する評価に影響していると考えられる。
（LINEはマイナスが大きく、SNSは小さい。ネット通販はマイナスが大きく、メルカリなどはむしろプラス）。
- TikTokやメルカリ、ラクマなどがプラスなのは、**年齢層**も関係している可能性。

データ利活用への人々の評価

◆ マクロ便益推計モデル

- 以下のモデルによって日本全国の年間評価額を推計する。

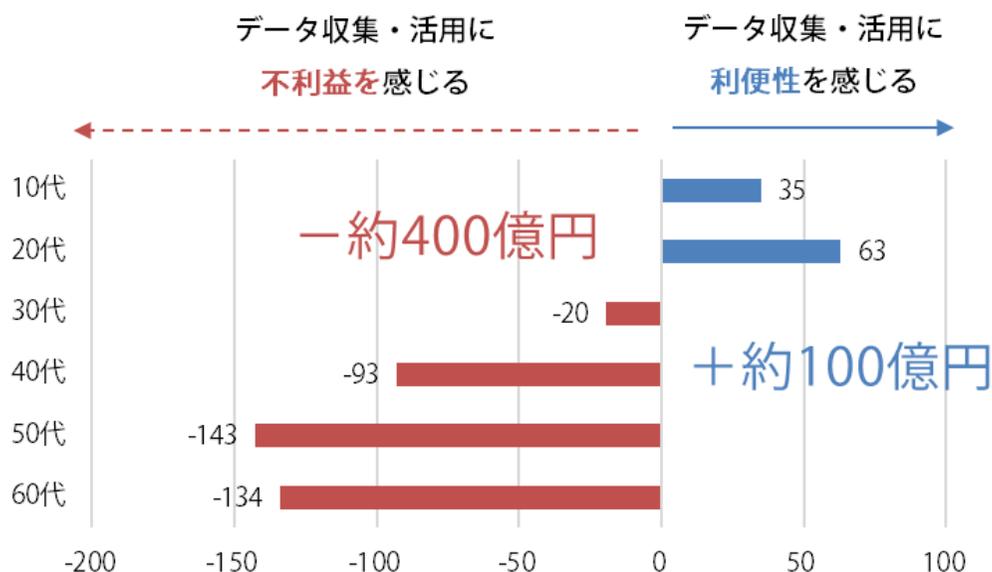
$$U_j = \sum_{k=1}^6 (\overline{Pay}_{jk} * 12 * Population_k * IRatio_k * SRatio_{jk})$$

- U_j : サービスjにおけるデータ利活用に対しての、人々の主観的な便益評価額の合計（年間）。
- \overline{Pay}_{jk} : サービスjのデータ利活用に対する、年代kの支払い意思額平均値（月間）。
- $Population_k$: 年代kの人口。
- $IRatio_k$: 年代kのインターネット利用率。
- $SRatio_{jk}$: 年代kのインターネットユーザに対する、サービスjの利用率。

データ利活用への人々の評価

◆ 便益のマクロ的な評価額の推計結果

- 10代と20代では**約100億円の便益**を得ていると評価している（ただし、10代は15～19歳に限る）。
- 30代以上では**約400億円の不利益**が発生しており、特に50代、60代で高い。
- 15～69歳が得ている便益評価額は合計**約－300億円**。



データ利活用への便益評価額（億円／年間）

- **子供のころ**からインターネットを利用している若年層と、新しいツールとして**大人になってから**導入された中高年以上で、認識が異なる可能性。
- これは、**ネット精通度**が高まれば支払い意思額が増加することとも整合性がとれる（後述の分析）。
- Milne & Boza（1999）、Dommeyer & Gross（2003）：プライバシー懸念水準に年齢差があり、**高年齢層は情報開示に慎重**。

データ利活用への人々の評価



21歳男性

個人データを取られることについては**困っておらず**、その理由として企業を信頼していること、万一データが漏れても自分にとって困ることがないと考えている。個人に関する情報が収集されていても、それによって**自分に適したものが表示されるならいい**と思っている。

いいなと思ったものもあれば、別にその機能自体要らないかな、と思うものもある。そのようにデータが収集されていて**気持ち悪いとは別段思わない**。仮にデータが取れない代わりに**有料**になるということになったら、LINEは有料でも使うが、**他のソーシャルメディアは使わない**と思う。



22歳女性

データが収集され活用されることは、**個人的にはOK**だと考えている。個人情報収集されることへの違和感も多少あるが、逆に、自分が落としてしまった製品や、気づかなかったことなど、**ミスを拾ってくれて助かる**と感じる。



28歳男性

評価は何が決定しているのか

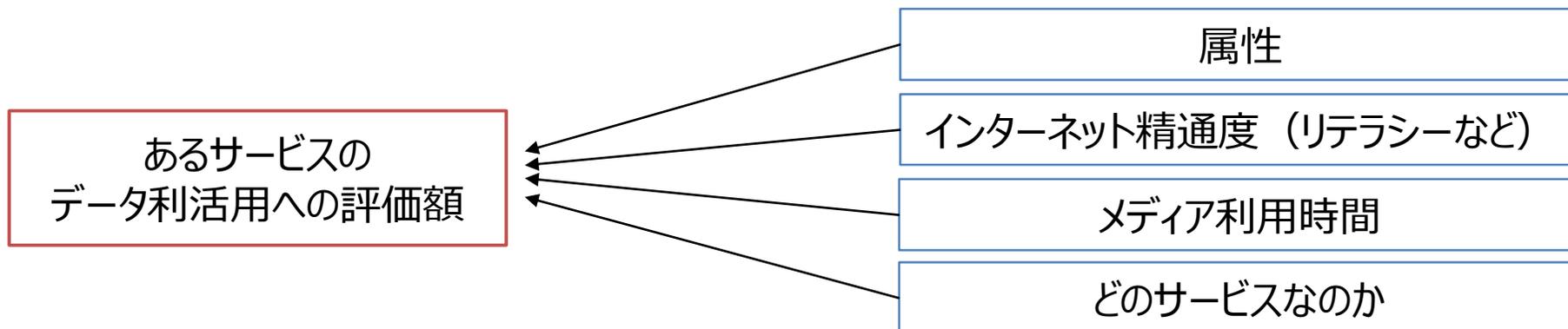
評価は何が決定しているのか

◆ 支払い意思額決定要因モデルでの分析

Pay_{ij}

$$= \alpha + \beta_1 Characteristics_i + \beta_2 Internet_i + \beta_3 Media_i + \gamma Service_j + \varepsilon_{ij}$$

- Pay_{ij} : 個人iのサービスjに対する支払い意思額。
- $Characteristics_i$: 個人iの属性ベクトル。性別、年齢、世帯年収、学歴、結婚有無。
- $Internet_i$: 個人iのインターネット精通度ベクトル。ネットリテラシー指標、ネット総利用時間、データ利活用認知指標。
- $Media_i$: 個人iのメディア利用時間ベクトル。ネットニュース閲読時間、新聞閲読有無、テレビ視聴時間、ラジオ聴取有無。
- $Service_j$: サービスjの固有効果ベクトル。
- α : 定数項。 β_1 、 β_2 、 β_3 : 各ベクトルにかかっているパラメータ。 ε_{ij} : 誤差項。



評価は何が決定しているのか

◆ 推定結果

項目	変数	係数	標準化係数	p値
属性	性別（男性）	5.043	0.027	0.00 **
	年齢	-0.537	-0.083	0.00 **
	世帯年収（万円）	-0.005	-0.019	0.00 **
	学歴（大卒）	-4.642	-0.024	0.00 **
	結婚（既婚）	8.376	0.044	0.00 **
ネット	ネットリテラシー指標	3.514	0.024	0.00 **
	ネット総利用時間（時間）	1.283	0.049	0.00 **
精通度	データ収集・活用認知指標	3.790	0.030	0.00 **
	ネットニュース閲読時間（時間）	4.108	0.024	0.00 **
メディア 利用時間	テレビ視聴時間（時間）	1.294	0.022	0.00 **
	新聞閲読	3.723	0.020	0.00 **
SNS	ラジオ聴取	2.511	0.013	0.01 **
	Facebook	5.772	0.013	0.00 **
	Twitter	5.456	0.013	0.00 **
	Instagram	6.803	0.015	0.00 **
	TikTok	11.963	0.022	0.00 **
メッセージアプリ	Messenger	10.164	0.019	0.00 **
ネットニュース	Yahoo! ニュース	1.635	0.004	0.00 **
	LINE NEWS	6.841	0.015	0.00 **
検索エンジン	Google	0.604	0.002	0.00 **
	Yahoo!	2.547	0.007	0.00 **
クチコミサイト	食べログ	6.002	0.014	0.00 **
	アットコスメ	9.526	0.019	0.00 **
	楽天トラベル	8.362	0.018	0.00 **
動画サイト	YouTube	3.155	0.009	0.00 **
	ニコニコ動画	9.764	0.020	0.00 **
フリマアプリ・オークションサイト	メルカリ	9.792	0.020	0.00 **
	ラクマ	10.216	0.020	0.00 **
ショッピングサイト	ヤフオク!	9.344	0.020	0.00 **
	Amazon	4.309	0.011	0.00 **
	楽天市場	3.190	0.008	0.00 **
	定数項	-22.594		0.00 **
	プーリング・F検定	4.970		0.00 **
	Hausman検定	80.490		0.00 **
サンプルサイズ			69949	

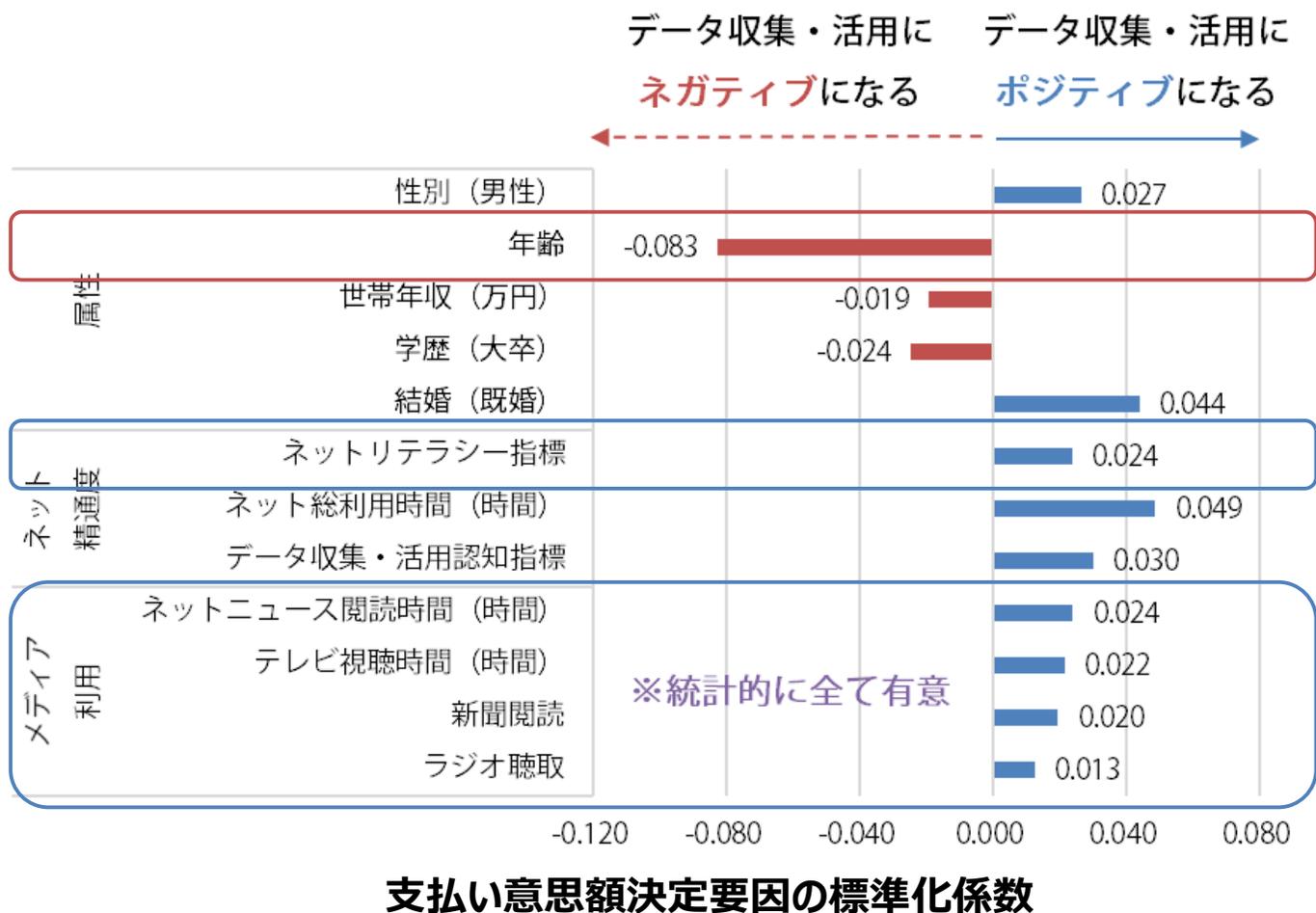
注1: **p<0.05, *p<0.1

注2: p値は不均一分散とクラスターに頑健な標準誤差から算出している。

- プーリング・F検定の結果、固定効果法が採択。
- Hausman検定の結果、固定効果法が採択。
- 全ての変数が有意水準1%で有意。

評価は何が決定しているのか

◆ 推定結果の概要



- 最も大きな影響を与えているのは**年齢**。
- **ネットリテラシー指標が高いとポジティブ**になる。利便性を認識するようになり、かつ、漠然とした不安感が減るため。
- **メディア利用時間が長いとポジティブ**になる。

考察・政策的含意

考察・政策的含意

◆ Key findings

- 90%程度の人は利便性と不安感が打ち消しあっており支払い意思額が0円。
- データ利活用がもたらす便益の総主観的評価は、日本全国で年間約－300億円。10代～20代で約＋100億円、30代以上で約－400億円。
- データ利活用への評価に年齢が最も大きな影響を与えていてマイナス。
- ネットリテラシーは、支払い意思額に有意に正。

政策的含意①：多様な価値観に配慮した制度設計をする

- データ利活用に対する評価は多様。とりわけこれから情報社会の中心となる10代・20代ではポジティブに受け止められている。政策決定者世代とのギャップ。
- 一律のデータ規制は将来の社会的厚生を下げる可能性がある。また、それによってサービス有料化などにつながった場合、被害は大きい。
- 一方で、悪用・明らかに想定されていない利用・個人情報保護の観点から明らかに問題のある利用などには断固とした態度で臨む。

考察・政策的含意

政策的含意②：利便性と不利益双方を考慮して制度設計する

- 「不安か」と聞かれると不安な人が多いが、利便性も評価している。
- **社会的厚生を最大**にするには、**双方をさらに調査**したうえで、エビデンスベースで最適な含意を導いていく必要がある。

政策的含意③：人々が適切なサービスを選択できる環境を作る

- 全体で約－300億円ということは、広義での**プライバシー・パラドックス**の存在を示唆。
- プラットフォーマーはデータ利活用について**オプションを用意**し、人々が自分の価値観に沿ったデータ利活用レベルを選択できるようにした方が良い。
- さらに、「**適切な選択**」のために、読みやすい**規約**を工夫する、ネットリテラシー教育を強化するなど、認知度を向上させる。

<プライバシー・パラドックス>

- ✓ プライバシー不安とインターネット上の自己開示行動の間に見られる乖離現象のこと（Baek, 2014; Taddicken, 2014）。
- ✓ 利便性を不安感が上回っているにもかかわらずデータを提供している人は、それに該当すると考えられる（三上, 2015）。

ご清聴ありがとうございました